

STIC Biotechnology Systems Branch

RAW SEQUENCE LISTING

ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/589, 255
Source: IFW
Date Processed by STIC: 04/04/2007

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 4.4.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):
U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/10/06

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER:

10/589,255

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped
 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000
- 9 Use of n's or Xaa's
 (NEW RULES) . Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 Invalid <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence. (see item 11 below)
- 11 ✓ Use of <220> Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown."
 Please explain source of genetic material in <220> to <223> section or use "chemically synthesized" as explanation. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32), also Sec. 1.823 of Sequence Rules
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n/Xaa "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid



IFWO

RAW SEQUENCE LISTING

DATE: 04/05/2007

PATENT APPLICATION: US/10/589,255

TIME: 07:48:16

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\04052007\J589255.raw

3 <110> APPLICANT: Japan Science and Technology Agency
 5 <120> TITLE OF INVENTION: Probe for detection of nuclear receptor agonist/antagonist
 and method for
 6 screening agonist and antagonist using the same
 8 <130> FILE REFERENCE: 04F055PCT
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/589,255
 C--> 10 <141> CURRENT FILING DATE: 2007-01-18
 10 <150> PRIOR APPLICATION NUMBER: JP 2004-35678
 11 <151> PRIOR FILING DATE: 2004-02-12
 13 <160> NUMBER OF SEQ ID NOS: 11
 15 <170> SOFTWARE: PatentIn version 3.1
 17 <210> SEQ ID NO: 1
 18 <211> LENGTH: 5
 19 <212> TYPE: PRT
 20 <213> ORGANISM: Unknown
 22 <220> FEATURE:
 23 <223> OTHER INFORMATION: any amino acid
 25 <220> FEATURE:
 26 <221> NAME/KEY: MISC_FEATURE
 27 <222> LOCATION: (2)..(3)
 28 <223> OTHER INFORMATION: any amino acid
 31 <400> SEQUENCE: 1
 W--> 33 Leu Xaa Xaa Leu Leu
 34 1 5
 37 <210> SEQ ID NO: 2
 38 <211> LENGTH: 11
 39 <212> TYPE: PRT
 40 <213> ORGANISM: Homo sapiens
 42 <300> PUBLICATION INFORMATION:
 43 <308> DATABASE ACCESSION NO: GeneBank/NP_671766
 44 <309> DATABASE ENTRY DATE: 2003-12-22
 45 <313> RELEVANT RESIDUES: (687)..(697)
 47 <400> SEQUENCE: 2
 49 His Lys Ile Leu His Arg Leu Leu Gln Glu Gly
 50 1 5 10
 53 <210> SEQ ID NO: 3
 54 <211> LENGTH: 239
 55 <212> TYPE: PRT
 56 <213> ORGANISM: Artificial
 58 <220> FEATURE:
 59 <223> OTHER INFORMATION: synthesized polypeptide
 61 <300> PUBLICATION INFORMATION:
 62 <308> DATABASE ACCESSION NO: GeneBank/GAG11884
 63 <309> DATABASE ENTRY DATE: 2003-05-21

Does Not Comply
 Corrected Diskette Needed
 Cpg-1,2

g/ 22137 Responses
 is Unknown or Artificial,
 pls explain the Source
 of genetic material.
 see Item 11 on Error
 summary sheet.

RAW SEQUENCE LISTING

DATE: 04/05/2007

PATENT APPLICATION: US/10/589,255

TIME: 07:48:16

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\04052007\J589255.raw

64 <313> RELEVANT RESIDUES: (1)..(238)

66 <400> SEQUENCE: 3

```

68 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
69 1          5          10          15
72 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
73          20          25          30
76 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
77          35          40          45
80 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
81          50          55          60
84 Leu Thr Trp Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
85 65          70          75          80
88 Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
89          85          90          95
92 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
93          100          105          110
96 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
97          115          120          125
100 Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
101          130          135          140
104 Asn Tyr Ile Ser His Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys Asn
105 145          150          155          160
108 Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
109          165          170          175
112 Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
113          180          185          190
116 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu
117          195          200          205
120 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
121          210          215          220
124 Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
125 225          230          235

```

128 <210> SEQ ID NO: 4

129 <211> LENGTH: 248

130 <212> TYPE: PRT

131 <213> ORGANISM: unknown

133 <220> FEATURE:

134 <223> OTHER INFORMATION: Artificial Sequence

136 <300> PUBLICATION INFORMATION:

137 <308> DATABASE ACCESSION NO: GeneBank/AAQ93355

138 <309> DATABASE ENTRY DATE: 2003-10-12

139 <313> RELEVANT RESIDUES: (1)..(238)

141 <400> SEQUENCE: 4

```

143 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
144 1          5          10          15
147 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
148          20          25          30
151 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
152          35          40          45

```

Same Error

RAW SEQUENCE LISTING

DATE: 04/05/2007

PATENT APPLICATION: US/10/589,255

TIME: 07:48:16

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\04052007\J589255.raw

```

155 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
156      50                      55                      60
159 Phe Gly Tyr Gly Leu Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys
160 65                      70                      75                      80
163 Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
164                      85                      90                      95
167 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
168                      100                     105                     110
171 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
172                      115                     120                     125
175 Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
176                      130                     135                     140
179 Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
180 145                      150                      155                      160
183 Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
184                      165                      170                      175
187 Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
188                      180                      185                      190
191 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu
192                      195                      200                      205
195 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
196                      210                      215                      220
199 Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser
200 225                      230                      235                      240
203 Gly Leu Arg Ser Thr Gly Ser Arg
204                      245
207 <210> SEQ ID NO: 5
208 <211> LENGTH: 246
209 <212> TYPE: PRT
210 <213> ORGANISM: Homo sapiens
212 <220> FEATURE:
213 <223> OTHER INFORMATION: part of a eucaryotic protein
215 <300> PUBLICATION INFORMATION:
216 <308> DATABASE ACCESSION NO: GeneBank/NP_036821
217 <309> DATABASE ENTRY DATE: 2004-01-23
218 <313> RELEVANT RESIDUES: (305)..(550)
220 <400> SEQUENCE: 5
222 His Thr Lys Lys Asn Ser Pro Ala Leu Ser Leu Thr Ala Asp Gln Met
223 1                      5                      10                      15
226 Val Ser Ala Leu Leu Asp Ala Glu Pro Pro Leu Ile Tyr Ser Glu Tyr
227                      20                      25                      30
230 Asp Pro Ser Arg Pro Phe Ser Glu Ala Ser Met Met Gly Leu Leu Thr
231                      35                      40                      45
234 Asn Leu Ala Asp Arg Glu Leu Val His Met Ile Asn Trp Ala Lys Arg
235                      50                      55                      60
238 Val Pro Gly Phe Gly Asp Leu Asn Leu His Asp Gln Val His Leu Leu
239 65                      70                      75                      80
242 Glu Cys Ala Trp Leu Glu Ile Leu Met Ile Gly Leu Val Trp Arg Ser
243                      85                      90                      95

```

RAW SEQUENCE LISTING

DATE: 04/05/2007

PATENT APPLICATION: US/10/589,255

TIME: 07:48:16

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\04052007\J589255.raw

```

246 Met Glu His Pro Gly Lys Leu Leu Phe Ala Pro Asn Leu Leu Leu Asp
247          100          105          110
250 Arg Asn Gln Gly Lys Cys Val Glu Gly Met Val Glu Ile Phe Asp Met
251          115          120          125
254 Leu Leu Ala Thr Ser Ser Arg Phe Arg Met Met Asn Leu Gln Gly Glu
255          130          135          140
258 Glu Phe Val Cys Leu Lys Ser Ile Ile Leu Leu Asn Ser Gly Val Tyr
259 145          150          155          160
262 Thr Phe Leu Ser Ser Thr Leu Lys Ser Leu Glu Glu Lys Asp His Ile
263          165          170          175
266 His Arg Val Leu Asp Lys Ile Asn Asp Thr Leu Ile His Leu Met Ala
267          180          185          190
270 Lys Ala Gly Leu Thr Leu Gln Gln Gln His Arg Arg Leu Ala Gln Leu
271          195          200          205
274 Leu Leu Ile Leu Ser His Ile Arg His Met Ser Asn Lys Gly Met Glu
275          210          215          220
278 His Leu Tyr Asn Met Lys Cys Lys Asn Val Val Pro Leu Tyr Asp Leu
279 225          230          235          240
282 Leu Leu Glu Met Leu Asp
283          245
286 <210> SEQ ID NO: 6
287 <211> LENGTH: 5
288 <212> TYPE: PRT
289 <213> ORGANISM: Artificial
291 <220> FEATURE:
292 <223> OTHER INFORMATION: synthesized oligopeptide
294 <400> SEQUENCE: 6
296 Gly Gly Asn Gly Gly
297 1          5
300 <210> SEQ ID NO: 7
301 <211> LENGTH: 11
302 <212> TYPE: PRT
303 <213> ORGANISM: Artificial
305 <220> FEATURE:
306 <223> OTHER INFORMATION: synthesized oligopeptide
308 <400> SEQUENCE: 7
310 His Lys Ile Ala His Arg Ala Ala Gln Glu Gly
311 1          5          10
314 <210> SEQ ID NO: 8
315 <211> LENGTH: 271
316 <212> TYPE: PRT
317 <213> ORGANISM: Homo sapiens
319 <220> FEATURE:
320 <223> OTHER INFORMATION: part of a eucaryotic protein
322 <300> PUBLICATION INFORMATION:
323 <308> DATABASE ACCESSION NO: GenBank/NM_015869
324 <309> DATABASE ENTRY DATE: 1996-11-04
325 <313> RELEVANT RESIDUES: (235)..(505)
327 <400> SEQUENCE: 8

```

RAW SEQUENCE LISTING

DATE: 04/05/2007

PATENT APPLICATION: US/10/589,255

TIME: 07:48:16

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\04052007\J589255.raw

```

329 Glu Ser Ala Asp Leu Arg Ala Leu Ala Lys His Leu Tyr Asp Ser Tyr
330 1          5          10          15
333 Ile Lys Ser Phe Pro Leu Thr Lys Ala Lys Ala Arg Ala Ile Leu Thr
334          20          25          30
337 Gly Lys Thr Thr Asp Lys Ser Pro Phe Val Ile Tyr Asp Met Asn Ser
338          35          40          45
341 Leu Met Met Gly Glu Asp Lys Ile Lys Phe Lys His Ile Thr Pro Leu
342          50          55          60
345 Gln Glu Gln Ser Lys Glu Val Ala Ile Arg Ile Phe Gln Gly Cys Gln
346 65          70          75          80
349 Phe Arg Ser Val Glu Ala Val Gln Glu Ile Thr Glu Tyr Ala Lys Ser
350          85          90          95
353 Ile Pro Gly Phe Val Asn Leu Asp Leu Asn Asp Gln Val Thr Leu Leu
354          100         105         110
357 Lys Tyr Gly Val His Glu Ile Ile Tyr Thr Met Leu Ala Ser Leu Met
358          115         120         125
361 Asn Lys Asp Gly Val Leu Ile Ser Glu Gly Gln Gly Phe Met Thr Arg
362          130         135         140
365 Glu Phe Leu Lys Ser Leu Arg Lys Pro Phe Gly Asp Phe Met Glu Pro
366 145         150         155         160
369 Lys Phe Glu Phe Ala Val Lys Phe Asn Ala Leu Glu Leu Asp Asp Ser
370          165         170         175
373 Asp Leu Ala Ile Phe Ile Ala Val Ile Ile Leu Ser Gly Asp Arg Pro
374          180         185         190
377 Gly Leu Leu Asn Val Lys Pro Ile Glu Asp Ile Gln Asp Asn Leu Leu
378          195         200         205
381 Gln Ala Leu Glu Leu Gln Leu Lys Leu Asn His Pro Glu Ser Ser Gln
382          210         215         220
385 Leu Phe Ala Lys Leu Leu Gln Lys Met Thr Asp Leu Arg Gln Ile Val
386 225         230         235         240
389 Thr Glu His Val Gln Leu Leu Gln Val Ile Lys Lys Thr Glu Thr Asp
390          245         250         255
393 Met Ser Leu His Pro Leu Leu Gln Glu Ile Tyr Lys Asp Leu Tyr
394          260         265         270
397 <210> SEQ ID NO: 9
398 <211> LENGTH: 30
399 <212> TYPE: PRT
400 <213> ORGANISM: Artificial
402 <220> FEATURE:
403 <223> OTHER INFORMATION: synthesized oligopeptide
405 <400> SEQUENCE: 9
407 Gly Gly Asn Gly Gly Gly Gly Asn Gly Gly Gly Gly Asn Gly Gly Gly
408 1          5          10          15
410 Gly Asn Gly Gly Gly Gly Asn Gly Gly Gly Gly Asn Gly Gly
411          20          25          30
414 <210> SEQ ID NO: 10
415 <211> LENGTH: 239
416 <212> TYPE: PRT
417 <213> ORGANISM: Homo sapiens

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/589,255

DATE: 04/05/2007
TIME: 07:48:17

Input Set : A:\pto.da.txt
Output Set: N:\CRF4\04052007\J589255.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 2,3

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 5

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:3,6,7,9,11

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/589,255

DATE: 04/05/2007

TIME: 07:48:17

Input Set : A:\pto.da.txt

Output Set: N:\CRF4\04052007\J589255.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:33 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:0